Available online at www.sciencedirect.com
SCIENCE DIRECT®
Molecular Phylogenetics and Evolution 34 (2005) 625–644
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Molecular systematics of the butterfly genus *Ithomia* (Lepidoptera: Ithomiinae): a composite phylogenetic hypothesis based on seven genes

Ricardo Mallarino[a], Eldredge Bermingham[a], Keith R. Willmott[b], Alaine Whinnett[b,c], Chris D. Jiggins[d,*]

[a] *Smithsonian Tropical Research Institute, Apartado 2072, Balboa, Panama*
[b] *The Natural History Museum, Cromwell Road, London, SW7 5BD, UK*
[c] *Galton Laboratory, University College London, 4 Stephenson Way, London NW1 2HE, UK*
[d] *Institute of Evolutionary Biology, University of Edinburgh, Edinburgh EH9 3JT, UK*

Received 13 August 2004; revised 25 October 2004
Available online 16 December 2004

Abstract

Butterflies in the nymphalid subfamily Ithomiinae are brightly colored and involved in mimicry. Here we present a phylogenetic hypothesis for 23 of the 24 species in the genus *Ithomia*, based on seven different gene regions, representing 5 linkage groups and 4469 bp. We sequenced varying length regions of the following genes: (1) elongation factor 1α (*Ef1α*; 1028 bp); (2) tektin (*tektin*; 715 bp); (3) wingless (*wg*; 405 bp); (4) ribosomal protein L5 (*RpL5*; 722 bp, exons 1, 2, 3, and introns 1 and 2); and (5) mitochondrial cytochrome oxidase I, II (*Co1* and *Co2* and intervening leucine tRNA; 1599 bp). The results show incongruence between some genetic loci, although when alternate topologies are compared statistically it was generally true that one topology was supported by a majority of loci sampled. This highlights the need to sample widely across the genome in order to obtain a well-supported phylogenetic hypothesis. A combined evidence topology is presented based on a Bayesian analysis of all the gene regions, except the fast-evolving *RpL5*. The resulting hypothesis is concordant with the most probable relationships determined from our topological comparisons, although in some parts of the tree relationships remain weakly supported. The tree suggests diversification has largely occurred within biogeographic regions such as Central America, the Amazon, the southern and northern Andes, with only occasional dispersal (or vicariance) between such regions. This phylogenetic hypothesis can now be used to investigate patterns of diversification across the genus, such as the potential role of color pattern changes in speciation.

Keywords: Lepidoptera; Mimicry; Phylogeny; Phylogenetic discordance; Speciation

1. Introduction

Phylogenetic approaches to the study of speciation are becoming increasingly common as molecular phylogenetic analyses of entire species clades become available (Barraclough and Vogler, 2000; Barraclough et al., 1998). These studies rely on the fact that species-level phylogenetic hypotheses permit both the rate and pattern of cladogenesis to be inferred. In combination with ecological and morphological information for the species concerned, phylogenies can be used to test hypotheses regarding which factors promote cladogenesis.

The wing patterns of butterflies are traits that play an important role in diversification. A number of groups show great diversity and rapid divergence in color

* Corresponding author. Fax: +44 131 650 6564.
E-mail address: chris.jiggins@ed.ac.uk (C.D. Jiggins).

1055-7903/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2004.10.021

pattern between both populations and species (Bates, 1862; Vane-Wright, 1979). Furthermore, the dual role of wing patterns in signaling to both predators and potential conspecific mates means that they may commonly cause reproductive isolation between divergent populations. In the case of aposematic warning colors, switches in color pattern between related species can generate both pre- and post-mating reproductive isolation and thus contribute to speciation (Jiggins et al., 2001, 2004a). Therefore, color patterns are traits likely to play a role in speciation, which are also easily recorded and therefore amenable to comparative analysis.

The nymphalid sub-family Ithomiinae is a diverse group of neotropical butterflies (around 360 species, Lamas, 2004). All are thought to be distasteful to predators and form part of diverse Müllerian mimicry rings that include heliconiine nymphalid butterflies and many other Lepidoptera, including day-flying moths (Beccaloni, 1997; Brown, 1984). The larvae feed primarily on hosts in the family Solanaceae, a family known for its defensive chemicals (Drummond and Brown, 1987). The genus *Ithomia* consists of 24 recognized species of relatively abundant and widespread butterflies. These are often involved in mimicry with other ithomiine species, notably in the genera *Napeogenes*, *Hypothyris*, *Scada*, and others. Males of this genus are readily recognized in the field by a distinctive bump on the ventral hindwing opposite the dorsal androconial hair pencils, which contains densely packed elongate androconial scales (Fox, 1940, 1968). This unusual morphological feature occurs in *Ithomia* and its close relative *Pagyris*, while several other morphological synapomorphies support the monophyly of *Ithomia* (Willmott and Freitas, in preparation). Due to the abundance of individuals, and hence relative ease of obtaining specimens, the strong morphological support for monophyly of the genus, and the diversity of mimetic relationships found within the group we chose the *Ithomia* as a test case for future species-level molecular phylogenetic analyses across the Ithomiinae. Understanding the phylogenetic relationships among members of this genus will hopefully provide insights into the biological processes underlying the diversification of the Ithomiinae.

The first step towards testing ecological hypotheses is the construction of a robust and reliable phylogenetic hypothesis. Gene sequence data provide an excellent source of characters that are likely to be independent of the morphological and ecological characters under study. In addition, the mode of transmission and genetic basis of sequence data can be determined explicitly, and they facilitate estimation of evolutionary rates (Avise, 1994). Nonetheless, it is becoming clear that all genomes are mosaics of regions with slightly different genealogical histories (Lerat et al., 2003; Pääbo, 2003; Rokas et al., 2003). Two recent studies have used whole genome sequences to investigate this phenomenon (Lerat et al., 2003; Rokas et al., 2003). Both showed that although there is considerable evidence for incongruence in phylogenetic signal between loci, nonetheless a consistent consensus topology was obtained from the majority of genes. Such studies emphasize the need to sample widely across the genome in reconstructing phylogenetic hypotheses (Rokas et al., 2003).

We here use seven different gene regions in order to infer a phylogenetic hypothesis for the genus *Ithomia*. The results show incongruence between some genetic loci and highlight the need for such a wide sampling of genomic regions. There are some areas in the resulting *Ithomia* phylogeny where congruence across many genes provides high confidence in the hypothesized relationships, and other areas where there is disagreement between loci suggesting additional analysis may be required. Nonetheless, a most probable pattern of relationships is presented that is supported by a majority of the loci sampled.

2. Materials and methods

2.1. Loci analyzed

Traditionally there has been an emphasis on the utility of mtDNA for inferring relationships among closely related species and species groups, with many examples in butterflies (Beltrán et al., 2002; Blum et al., 2003; Davies and Bermingham, 2002; Sperling and Harrison, 1994). For comparison, we used the same mtDNA loci studied previously. In addition, we include three nuclear loci that have been shown to provide phylogenetically useful information in recent work on butterflies and other insects, *Ef1α*, *tektin*, and *wingless* (Brower and DeSalle, 1998; Cho et al., 1995; Whinnett et al., in prep.). We also developed a gene region that has not been previously reported in phylogenetic studies, a fragment of the *Ribosomal Protein L5* that contains two highly variable intron regions. This locus is useful for resolving relationships among closely related taxa. Our sampling therefore includes more loci than is common in phylogenetic studies.

2.2. Samples, DNA extraction, amplification, and sequencing

Thorax or leg samples of butterflies were either collected by the authors or obtained through donation (Table 1). We obtained a total of 23 out of the 24 species currently recognized in *Ithomia*. We were unable to obtain samples from *I. leila*. Butterflies in the closely related genera *Napeogenes* and *Pagyris* (Willmott and Freitas, in prep.) were used as outgroups. The taxonomy of the species analyzed follows Lamas (2004), except for two cases. First, *I. pseudoagalla*, formerly considered a

Table 1
Specimens of *Ithomia* and related genera included in the present study

Taxa	ID #	Locality	Lat	Long
Ithomia agnosia agnosia Hewitson	8901	Macas—Puyo KM70, Morona-Santiago, Ecuador	2.317 S	78.512 W
Ithomia agnosia agnosia Hewitson	8902	Macas—Puyo KM70, Morona-Santiago, Ecuador	2.317 S	78.512 W
Ithomia agnosia agnosia Hewitson	8895	Puente Aguas Verdes, Rioja, San Martín, Peru	5.685 S	77.658W
Ithomia agnosia agnosia Hewitson	8896	Chumía, San Martín, Peru	6.695 S	76.186 W
Ithomia amarilla Haensch	9105	Río Tiputini, Orellana, Ecuador	0.422 S	76.005 W
Ithomia arduinna arduinna d'Almeida	8942 (3468)	Río Madre de Dios, Cuzco, Peru	12.250 S	70.900 W
Ithomia avella epona Hewitson	9102	La Bonita-Tulcan rd, Sucumbíos, Ecuador	0.56 N	77.53 W
Ithomia avella epona Hewitson	9103	Río Sucio, Sucumbíos, Ecuador	0.48 N	77.56 W
Ithomia celemia lurida Haensch	9151	Municipio Ciudad Bolívar, Antioquia, Colombia	5.973 N	75.694 W
Ithomia cleora Hewitson	8905	Vilcabamba, Loja, Ecuador	4.25 S	79.25 W
Ithomia cleora Hewitson	8904	Vilcabamba, Loja, Ecuador	4.25 S	79.25 W
Ithomia diasia hippocrenis H. W. Bates	8286	Cerro Campana, Panama	8.687 N	79.920 W
Ithomia diasia hippocrenis H. W. Bates	8294	Santa Rita Ridge, Colón, Panama	9.32 N	79.78 W
Ithomia diasia hippocrenis H. W. Bates	2985	Continental Divide, Fortuna, Chiriquí, Panama	8.785 N	82.214 W
Ithomia drymo Hübner	9100 (B-16-5)	Serro do Japi Municipal Park, Jundiai, São Paulo, Brazil	23.143 S	46.867 W
Ithomia eleonora Haensch	9114	Qbda. Siete Jeringas, Junín, Peru	11.20 S	75.40 W
Ithomia eleonora Haensch	9117	Qbda. Siete Jeringas, Junín, Peru	11.20 S	75.40W
Ithomia ellara Hewitson	9115	Qbda. Siete Jeringas, Junín, Peru	11.20 S	75.40 W
Ithomia heraldica H. W. Bates	8005	Qbda. Hornito, Fortuna, Chiriquí, Panama	8.693 N	82.225 W
Ithomia heraldica H. W. Bates	8006	Qbda. Hornito, Fortuna, Chiriquí, Panama	8.693 N	82.225 W
Ithomia hyala hyala Hewitson	8910	Hotel Tinalandia, Pichincha, Ecuador	0.284 S	79.050 W
Ithomia hyala hyala Hewitson	8911	Hotel Tinalandia, Pichincha, Ecuador	0.284 S	79.050 W
Ithomia hyala ssp. nov.	8844	Cerro Pirre, Darién, Panama	7.764 N	77.721 W
Ithomia hymettia Staudinger	9150	Municipio Ciudad Bolívar, Antioquia, Colombia	5.812 N	76.013 W
Ithomia iphianassa panamensis H. W. Bates	8102	Achiote Road, Colón, Panama	9.218 N	79.979 W
Ithomia iphianassa panamensis H. W. Bates	8804	Cana, Darién, Panama	7.756 N	77.684 W
Ithomia iphianassa panamensis H. W. Bates	8805	Cana, Darién, Panama	7.756 N	77.684 W
Ithomia iphianassa ssp. nov.	8912	Hotel Tinalandia, Pichincha, Ecuador	0.284 S	79.050 W
Ithomia iphianassa ssp. nov.	8914	Hotel Tinalandia, Pichincha, Ecuador	0.284 S	79.050 W
Ithomia iphianassa ssp. nov.	8107	Pedro Vicente Maldonado, Ecuador	0.118 N	79.059 W
Ithomia jucunda centromaculata Weymer	8873	Cana, Darién, Panama	7.756 N	77.684 W
Ithomia lagusa peruana Salvin	8915	Tarapoto—Yurimaguas KM8, San Martín, Peru	6.501 N	76.366 W
Ithomia lagusa peruana Salvin	8916	Tarapoto—Yurimaguas KM8, San Martín, Peru	6.501 N	76.366 W
Ithomia lagusa theuda Hewitson	9096 (E-30-6)	Quebrada Chorillos, Zamora-Chinchipe, Ecuador	4.069 S	78.957 W
Ithomia lichyi dalmeidai R. M. Fox	8931 (3530)	Río Inambari, Madre de Dios, Peru	12.600 S	70.083 W
Ithomia lichyi dalmeidai R. M. Fox	8933 (3534)	Río Inambari, Madre de Dios, Peru	12.600 S	70.083 W
Ithomia lichyi dalmeidai R. M. Fox	8937 (3538)	Río Inambari, Madre de Dios, Peru	12.600 S	70.083 W
Ithomia patilla Hewitson	8028	Finca Café Duran, Chiriquí, Panama	8.839 N	82.715 W
Ithomia patilla Hewitson	8031	Finca Café Duran, Chiriquí, Panama	8.839 N	82.715 W
Ithomia praeithomia Vitale & Bollino	9116	Qbda. Siete Jeringas, Junín, Peru	11.20 S	75.40 W
Ithomia pseudoagalla Rebel	8898	Hotel Tinalandia, Pichincha, Ecuador	0.284 S	79.050 W
Ithomia pseudoagalla Rebel	8899	Hotel Tinalandia, Pichincha, Ecuador	0.284 S	79.050 W
Ithomia salapia aquinia Hopffer	8918	Tarapoto—Yurimaguas KM30, San Martín, Peru	6.501 N	76.366 W
Ithomia salapia aquinia Hopffer	8919	Tarapoto—Yurimaguas KM30, San Martín, Peru	6.501 N	76.366 W
Ithomia salapia derasa Hewitson	8921	Puente Río Serranoyacu, Rioja, San Martín, Peru	5.675 S	77.675 W
Ithomia salapia derasa Hewitson	8922	Puente Río Serranoyacu, Rioja, San Martín, Peru	5.675 S	77.675 W
Ithomia salapia salapia Hewitson	8133	Jatun Sacha, Napo, Ecuador	0.782 S	77.804 W
Ithomia salapia salapia Hewitson	8135	Jatun Sacha, Napo, Ecuador	0.782 S	77.804 W
Ithomia terra vulcana Haensch	8041	Finca Café Duran, Chiriquí, Panama	8.839 N	82.715 W
Ithomia terra vulcana Haensch	8040	Finca Café Duran, Chiriquí, Panama	8.839 N	82.715 W
Ithomia terra vulcana Haensch	2990	Continental Divide, Fortuna, Chiriquí, Panama	8.785 N	82.214 W
Ithomia terra terrana Haensch	8924	Mindo, Pichincha, Ecuador	0.033 S	78.800 W
Ithomia terra terrana Haensch	8925	Mindo, Pichincha, Ecuador	0.033 S	78.800 W
Ithomia terra ssp. nov.	9113	Qbda. Siete Jeringas, Junín, Peru	11.20 S	75.40 W
Ithomia xenos H. W. Bates	8003	Qbda. Hornito, Fortuna, Chiriquí, Panama	8.693 N	82.225 W
Ithomia xenos H. W. Bates	8004	Qbda. Hornito, Fortuna, Chiriquí, Panama	8.693 N	82.225 W
Ithomia xenos H. W. Bates	8332	Qbda. Hornito, Fortuna, Chiriquí, Panama	8.693 N	82.225 W
Napeogenes cranto paedaretus Goldman & Salvin	8312	Qbda. Hornito, Fortuna, Chiriquí, Panama	8.693 N	82.225 W
Napeogenes cranto paedaretus Goldman & Salvin	8313	Qbda. Hornito, Fortuna, Chiriquí, Panama	8.693 N	82.225 W
Pagyris cymothoe sylvella Hewitson	8945	Hotel Tinalandia, Pichincha, Ecuador	0.284 S	79.050 W
Pagyris cymothoe sylvella Hewitson	8946	Hotel Tinalandia, Pichincha, Ecuador	0.284 S	79.050 W

Numbers in parentheses represent original codes of donated samples. Decimal points given for latitude and longitude reflect differing accuracy depending on whether GPS measurements were taken at the collecting site (3 dp) or determined from map coordinates (2 dp).

race of *I. agnosia*, is here ranked as a species based on the molecular data (see Section 4). Second, it has been suggested that *I. hymettia* might be a sub-species of *I. lagusa* (Lamas, 2004). However, the two forms are sympatric in Colombia (Sandra Muriel, Pers. Comm.) suggesting species status, a conclusion supported by molecular data (see Section 4). The wings of all sampled species are currently in the private collection of CD Jiggins and will be accessioned to a museum collection once morphological studies are complete.

Samples for DNA analysis were preserved either in a 20% Dimethylsulfoxide, 0.25 M EDTA and saturated NaCl solution or in 95% Ethanol and are stored in the Smithsonian Tropical Research Institute in Panama. From each individual, one-half thorax was ground in lysis buffer and genomic DNA was extracted using the Qiagen DNeasy TissueKit (Qiagen Incorporated), following the protocol for animal tissue extraction, and preserved at −80 °C.

We amplified a 1599 bp region of mtDNA spanning the 3′ end of *Co1*, *leucine tRNA*, and the entire *Co2* gene from individual genomic DNA using polymerase chain reaction (PCR) and published primers (Table 2). In addition, we amplified a 1028 bp fragment of the nuclear *elongation factor 1 alpha* gene (*Ef1α*) using primers modified from Cho et al. (1995), a 405 bp fragment of the *wingless* gene (*wg*) using published primers (Brower and DeSalle, 1998), and a 715 bp fragment of the *tektin* gene using primers designed by Whinnett et al. (in prep). Finally a 722 bp fragment of the Ribosomal Protein L5 (*RpL5*) gene was amplified using novel primers (Table 2). These were designed from a sequence obtained from a cDNA library prepared from wing tissue of *Heliconius erato* and identified as a homolog of the *RpL5* gene (GenBank Accession No. CO377790). This sequence was aligned to the corresponding *Bombyx mori* sequence obtained from the Silkbase EST database (Mita et al., 2003). Primers were located in regions showing a high degree of conservation between the two species, and were found to amplify the region in both heliconiine and ithomiine butterflies. Double stranded DNA for each of the gene regions was synthesized in 15 μl reactions containing 2 μl DNA, 1× PCR buffer, 1 mM $MgCl_2$, 0.8 mM dNTPs, 0.5 mM of each primer, and 0.75 U/μl *Taq* Qiagen. Amplification conditions for the different genes are shown in Table 2.

The PCR products were electrophoretically separated from unincorporated nucleotides and primers in 1.5% low-melting point agarose gels run in TAE (Tris–acetate low-EDTA buffer, pH 7.8) containing ethidium bromide (1 μg/ml). The amplification product was cut, the agarose was incubated at 70 °C for 10 min, digested with 1 μl GELase (Epicenter Technologies, Madison WI), followed by a 3 h incubation period at 45 °C.

One to two microliters (μl) of the purified PCR product was used as a template in a 10 μl cycle sequence reac-

Table 2
Amplification primers and conditions used to amplify the *Ithomia* genes used in this study

Gene region	Primer names (E, I)	Sequence (5′ to 3′)	Reference	PCR conditions			
				Denaturation	Annealing	Extension	No. of cycles
Co1	Jerry (E)	CAACATTTATTTTGATTTTTTGG	Simon et al. (1994)	94 °C for 45 s	48 °C for 45 s	72 C for 1 min	5
	Pat (E)	TCCAATGCACTAATCTGCCATATTA		94 °C for 45 s	52 °C for 45 s	72 °C for 1.5 min	30
	Dick (I)	CCAACAGGAATTAAAATTTTTAGATGATTAGC					
Co2	George (E)	TAGGATTAGCTGGAATACC					
	Phyllis (I)	GC(ACT)GGTAAAATAGTTCAAATTAATTC					
	Imelda (E)	CATTAGAAGTAATTGCTAATTTACTA					
Ef1α	EF1a-for (E)	CTGAGCGYGARCGTGGTAT	Cho et al. (1995)	94 °C for 30 s	55 °C for 1 min	72 °C for 1 min	35
	EF1a-rev (E)	ACAGCNACKGTYTGYCTCAT					
tektin	Tektin A (E)	ACC AGT GGR GAY ATY CTW GG	Whinnett et al. (unpublished)	94 °C for 1 min	55 °C for 1 min	72 °C for 1.5 min	35
	Tektin 3 (E)	CGC AGT TTY TGA TRC TYT					
wg	LepWG1 (E)	GARTGYAARTGYCAYGGYATGTCTGG	Brower and DeSalle (1998)	94 °C for 1 min	48 °C for 1 min	72 °C for 1 min	35
	LepWG2 (E)	ACTICGCRCACCARTGGAATGTRCA					
RpL5	RpL5-F44 (E)	TCCGACTTTCAAACAAGGATG	Jiggins et al. unpub.	94 for 1 min	55 °C for 1 min	72 °C for 1.5 min	35
	RpL5-R441 (E)	CATATCCTGGGAATCTCTTGATG					

References are given for previously published primers. E, external primer; I, internal primer.

tion containing 1 μl Big Dye 3.1, 1 μl primer, and distilled-deionized water (ddH_2O). Cycling conditions followed standard ABI protocols (see Table 2 for sequencing primers). The cycle sequencing product was diluted in 10 μl ddH_2O, purified in Centrisep columns filled with 700 μl G-50 Sephadex, and air dried. Each sample was resuspended in 10 μl Hi-Di formamide, denatured at 95 °C for 2 min and then run on an ABI 3100 Automated DNA Sequencer following the manufacturer's protocol.

Chromatograms were edited and aligned using Sequencher 4.1 (Gene Codes Corporation) and nucleotide sequences were visually inspected for miscalls, reading frame errors, and termination codons. Variable length intron sequences were aligned manually in Sequencher 4.1.

2.3. Molecular characterization and phylogenetic analysis

Nucleotide sequences for protein coding mtDNA and nDNA sequences were translated to functional peptide sequences using MacClade 4.0 (Maddison and Maddison, 1997) as a further check. Data partition homogeneity tests (Farris et al., 1995) implemented in PAUP* 4.0b4a (Swofford, 2000) were carried out to determine whether different data partitions (*Co1*, *leucine tRNA*, *Co2*, *Ef1α*, *tektin*, *wg*, and *RpL5*) were congruent and could be analyzed together. Data sets were then analyzed using MODELTEST 3.04 (Posada and Crandall, 1998) to determine the sequence evolution model that best described our data. The best fit model chosen by Modeltest using likelihood ratio tests was used to perform a Bayesian analysis using the program Mr. Bayes 3.0 (http://morphbank.ebc.uu.se/mrbayes/) (Huelsenbeck et al., 2001). We ran four simultaneous Markov chains for 1,000,000 generations starting from random initial trees, and sampled a tree every 100 generations. Data from the first 100,000 generations were discarded, after confirming that likelihood values had stabilized prior to the 100,000th generation. The consensus phylogeny and posterior probability of nodes were calculated from trees sampled following the 100,000th generation. This process was repeated three times independently for each data set to account for the possibility of multiple optima. Model parameters for the Bayesian consensus tree were calculated in PAUP*. For comparison, we also conducted Maximum Parsimony (MP) analyses in PAUP*, using a heuristic search with TBR branch swapping for each data partition separately and for the combined evidence data set. In each case the consensus tree was calculated using 50% majority rule, and confidence in branches was assessed by bootstrapping with TBR branch swapping (1000 replicates). For the combined evidence analysis, both majority rule and strict consensus trees were calculated for comparison. The program TreeRot (Sorenson, 1999) was used to calculate partitioned Bremer support values to assess the contribution of each gene to the overall Bremer support values.

2.4. Molecular clock analysis

We used PAUP* to calculate the $-\ln L$ values for non-clock and clock-enforced ML trees in order to test the constancy of evolution rates across the tree. The trees were compared using likelihood ratio tests.

3. Results

3.1. mtDNA

We sequenced 60 specimens representing 25 species, 23 from the *Ithomia* and two outgroup species, *Pagyris cymothoe* and *Napeogenes cranto*. The final mitochondrial alignment consisted of 1599 base pairs, representing 822 bp of the *Co1* gene, the entire 62 bp of the intervening *leucine tRNA*, and 715 bp of the *Co2* gene (GenBank Accession Nos. AY713036–AY713091). Of the 1599 nucleotide sites examined, 336 (21%) were variable and 284 (17%) were phylogenetically informative. The three individuals representing *I. avella* did not amplify for the entire mitochondrial fragment, and are only represented by the first 807 bp of the *Co1* gene (GenBank Accession Nos. AY713115–AY713117). These specimens were obtained as dried material and the lack of amplification was presumably due to sample degradation. There was no length variation between aligned regions of our mtDNA sequences.

Partition homogeneity tests showed the *Co1*, *leucine tRNA*, and *Co2* genes to be congruent with one another ($P = 0.13$), thus all subsequent analyses were carried out on the combined mtDNA data. The best-fit model chosen to describe the evolution of the mtDNA sequences was the GTR + I + G model (Yang, 1994) that incorporates six rate parameters, the proportion of invariable sites, and a gamma-distribution of rate heterogeneity (for parameter value estimates see Table 3). The Bayesian tree produced under this model was well resolved with virtually all species-level relationships well supported (Fig. 1). The mtDNA tree shows the species grouped into five higher clades supported by Bayesian posterior probabilities greater or equal than 0.95: *I. iphianassa* ($N = 5$ species), *I. hyala* ($N = 1$ species), *I. heraldica* ($N = 2$ species) *I. eleonora* ($N = 7$ species), and *I. agnosia* ($N = 8$ species).

3.2. Nuclear DNA

Four nuclear gene regions were sequenced: *Ef1α* ($N = 46$ individuals), *tektin* ($N = 42$ individuals), *wg* ($N = 44$ individuals), and *RpL5* ($N = 23$ individuals). In *Heliconius* it is known that *Ef1α*, *wg* and *RpL5* are on

Table 3
Parameter estimates of sequence evolution for the *Ithomia* genes used in this study

Gene region	Model	Base composition				Substitution rates								
								Tr		Tv				
		a	c	g	t	I	α	a–g	c–t	a–c	a–t	c–g	g–t	Tr/Tv
Co1 + Co2	GTR + I + G	0.3345	0.1062	0.1139	0.4455	0.674	0.8417	20.8742	62.7202	6.3276	3.9125	0	1	—
Ef1α	TrNef + G	0.25	0.25	0.25	0.25	0	0.0866	3.9134	9.8707	1	1	1	1	—
tektin	TrN + G	0.3654	0.1566	0.2333	0.2448	0	0.4019	2.75	5.21111	1	1	1	1	—
Wingless	K80 + G	0.25	0.25	0.25	0.25	0	0.3804	—	—	—	—	—	—	3.7497
Intron	TrN + G	0.3042	0.1464	0.2049	0.3444	0	0.2479	1.2807	4.1383	1	1	1	1	—

different chromosomes (Jiggins et al., 2004b; Tobler et al., 2004). Information is not available for *tektin*. Hence, although it is not known whether linkage relationships are conserved between *Heliconius* and *Ithomia*, it nonetheless seems unlikely that these loci are tightly linked in our study species. The number of individuals sequenced per locus varied for three reasons. First, as stated above, the DNA yielded by some of the dried material obtained through donation was degraded and led to amplification problems (*tektin* for *I. avella* and *wg* for *I. amarilla*). Second, the relative lack of genetic variation at some nuclear genes made it unnecessary to sequence all conspecific individuals. An initial survey performed with randomly selected species found that there was no sequence divergence among conspecific taxa in the *EF1α* gene, despite the inclusion of different races. Third, in the case of the *RpL5* intron, the locus was selected to address a specific set of phylogenetic relationships, and additional analysis across more distantly related *Ithomia* was considered impractical, mainly due to the difficulty in aligning intron sequences among distantly related taxa (Beltrán et al., 2002).

When the nuclear coding genes under investigation were evaluated under the partition homogeneity test, *tektin* and *Ef1α* showed significant heterogeneity ($P = 0.03$), whereas *wg* was not significantly incongruent when evaluated against *tektin* ($P = 0.57$) or *Ef1α* ($P = 0.81$). Similarly, the *RpL5* intron proved to be significantly heterogeneous when compared to *tektin* and *Ef1α* ($P = 0.01$), but not in comparison with *wg* ($P = 1.00$). Thus, we analyzed each nuclear gene independently and present the results for each gene in turn.

3.3. Ef1α

We obtained sequences for 46 individuals for the *Ef1α* gene (GenBank Accession Nos. AY704009–AY704054) corresponding to positions 2285–3312 in the *Drosophila* F1 reference sequence (GenBank Accession No. X06869). This region also corresponds to bases 19–1046 in the region used for a recent analysis of higher systematics in the Nymphalidae (Wahlberg et al., 2003), GenBank Accession No. AY090172. It is known that the *Ef1α* gene exists as two copies in some insects such as *Drosophila* and *Apis*, although there is no evidence that this is the case in the Lepidoptera (Danforth and Ji, 1998). Nonetheless, we aligned the amino acid sequences from one of our specimens (8102) to both the F1 and F2 copies of the *Drosophila* gene. Of the 44 amino acid residues that distinguish the two copies of the *Ef1α* gene in *Drosophila*, 12 grouped the *Ithomia* sequence with the F2 *Drosophila* copy, and 19 grouped it with the F1 copy of the gene. The *Ithomia Ef1α* that we sequenced is therefore intermediate between the two *Drosophila* copies. Following arguments presented elsewhere (Cho et al., 1995; Wahlberg et al., 2003), our data provide no evidence of duplicate and functionally distinct *Ef1α* copies in Lepidoptera, in contrast to the situation in *Apis* and *Drosophila*. Only two amino acid residues varied across the 44 *Ithomia* and outgroup individuals analyzed.

Of the 1028 *Ef1α* nucleotide sites analyzed, 118 (11%) were variable and 99 (9%) were phylogenetically informative. No sequences varied in length. Modeltest identified the TrNef + G model (Tamura and Nei, 1993) as the best fit to the *Ef1α* data (Table 3 provides model parameter estimates for all genes). The tree produced by Bayesian methods (Fig. 2) has well resolved relationships, particularly at deeper phylogenetic nodes, and shows some topological differences with respect to the mtDNA tree, discussed in detail below.

3.4. Tektin

We sequenced 42 individuals for the *tektin* gene (GenBank Accession Nos. AY704055–AY704096) corresponding to positions 350–1064 of the testis specific *Bombyx BmTST* (GenBank Accession No. AB056651). Of the 715 nucleotide sites examined, 131 (18%) were variable and 104 (14%) were phylogenetically informative. There was no length variation in any of the sequences. The best-fit model describing the data was TrN + G (Tamura and Nei, 1993). The Bayesian tree (Fig. 3) is well resolved, and has a well-supported topology nearly identical to the *Ef1α* gene tree. The only major difference between the *tektin* and *Ef1α* gene trees regards the relative placement of *I. hyala* (compare Figs. 2 and 3).

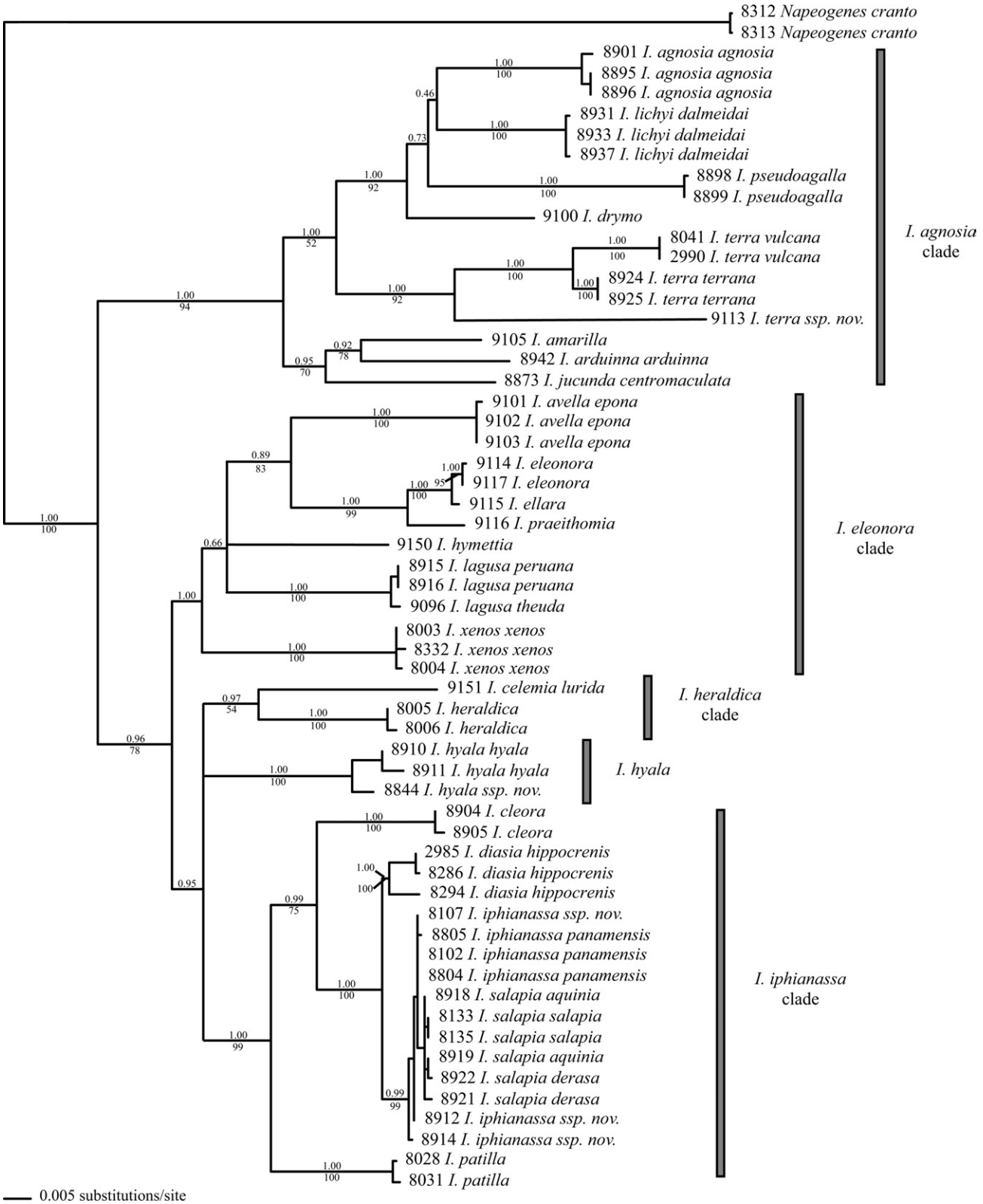

Fig. 1. Phylogenetic hypothesis for *Ithomia* based on 1599 nucleotide sites (336 variable) of mtDNA. Tree topology was inferred using Bayesian methods. Shown is the consensus phylogeny calculated from trees sampled following the 100,000th generation based on the GTR + I + G model. Bayesian probabilities (above) and parsimony bootstrap support (below) is given for each branch.

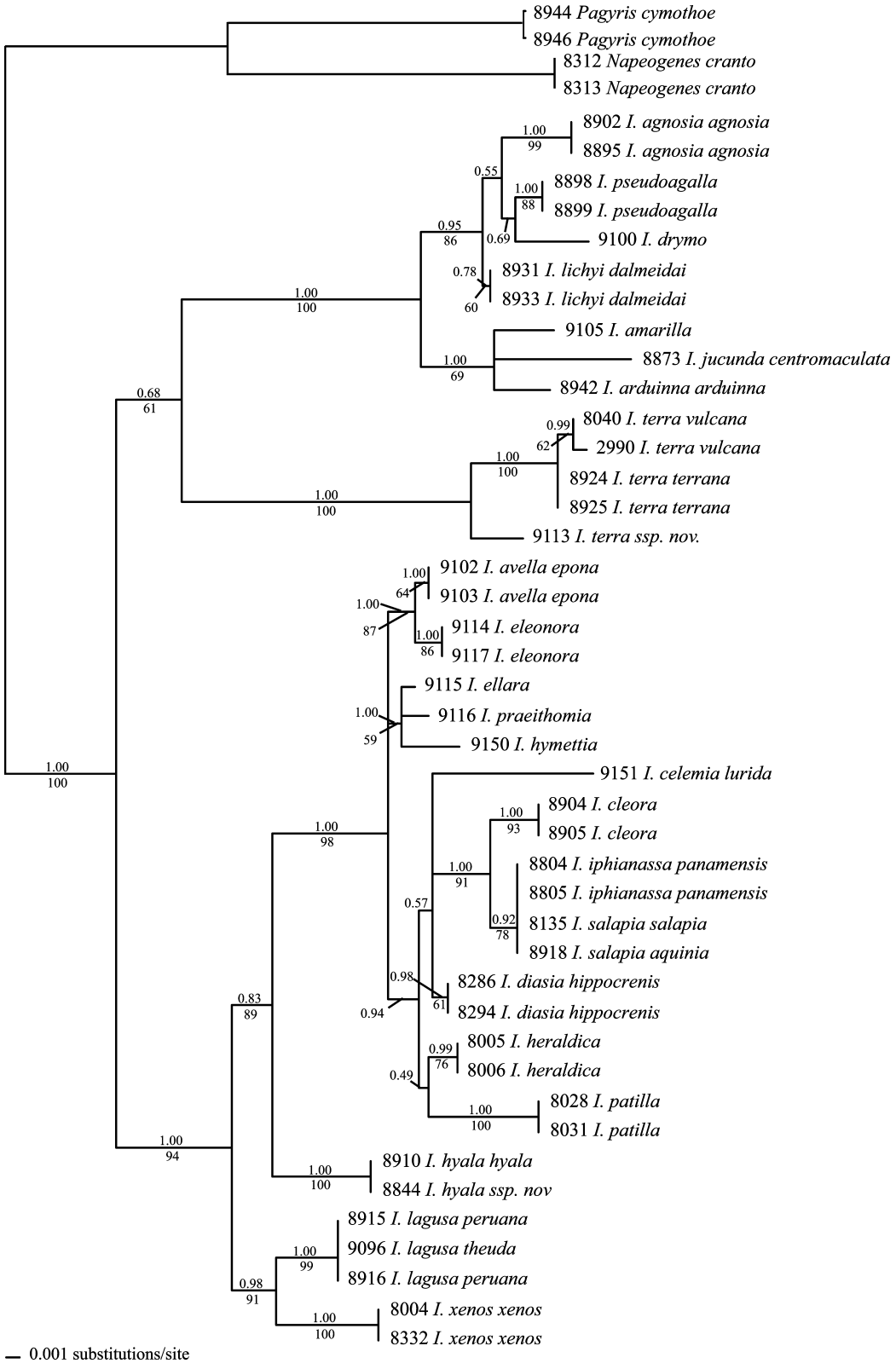

Fig. 2. Phylogenetic hypothesis for *Ithomia* based on 1028 nucleotide sites (118 variable) of the *Ef1*α gene. Tree topology was inferred using Bayesian methods. Shown is the consensus phylogeny calculated from trees sampled following the 100,000th generation based on the TrNef + G model (Tamura and Nei, 1993) model. Bayesian probabilities (above) and parsimony bootstrap support (below) is given for each branch.

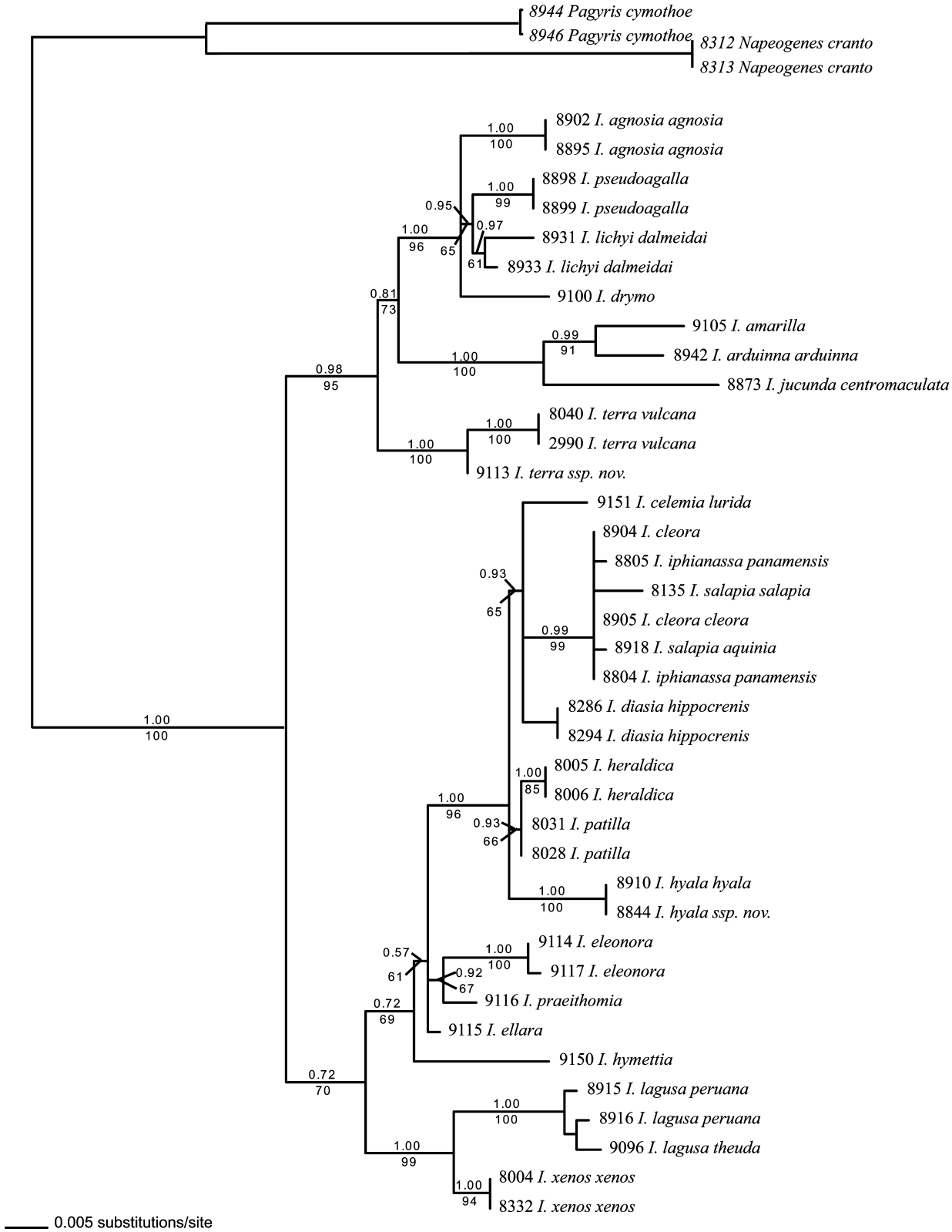

Fig. 3. Phylogenetic hypothesis for *Ithomia* based on 715 nucleotide sites (131 variable) of the *tektin* gene. Tree topology was inferred using Bayesian methods. Shown is the consensus phylogeny calculated from trees sampled following the 100,000th generation based on the TrN + G (Tamura and Nei, 1993) model. Bayesian probabilities (above) and parsimony bootstrap support (below) is given for each branch.

3.5. Wingless

A 405 bp region of the *wg* gene was sequenced for 44 *Ithomia* and outgroup individuals (GenBank Accession Nos. AY704097–AY704140). We analyzed 405 nucleotide sites corresponding to positions 279–683 in a reference sequence of *Precis coenia* (GenBank Accession No. L42142), of which 54 (13%) were variable and 43 (11%) were phylogenetically informative. No length variation was observed in any of the *wg* sequences analyzed. Modeltest indicated the K80 + G model (Kimura, 1980) was the best fit to the *wg* data. The Bayesian tree generated from the *wg* data (Fig. 4) was poorly resolved, but was generally consistent with *tektin* and *Ef1*α.

3.6. Rpl5 intron

To better determine phylogenetic relationships between *I. salapia* and *I. iphianassa* and close relatives, we analyzed introns 1 and 2 of the ribosomal protein L5 (*RpL5*) gene for 23 individuals (GenBank Accession Nos. AY713092–AY713114). The *Bombyx mori* EST sequence was used as a reference (GenBank Accession No. AF008229) to determine the nucleotide fragments representing exons and introns in our alignment. Bases 1–92 in our alignment represent the first exon, corresponding to bases 239–330 of the *Bombyx RpL5* sequence. The second exon, positions 292–451 in our alignment, corresponds to bases 331–490 in *Bombyx*, and the third exon fragment, positions 643–713 in our alignment, corresponds to positions 492–562 in *Bombyx*. It is not possible to compare intron positions between *Ithomia* and *Bombyx* given the lack of intron sequence registered in GenBank for *Bombyx*. For the *Ithomia* sequences analyzed the inferred length of intron 1 varied from 94 to 99 bp, and the length of intron 2 from 287 to 289 bp. Nonetheless, manual sequence alignments were generally unambiguous. Our alignment contained 722 bp, of which 88 (12%) were variable and 59 (8%) were phylogenetically informative. Unlike the other gene regions, our *RpL5* alignment also included several indels, which were mapped onto the tree subsequent to the ML analysis (Fig. 5). All *RpL5* data were acquired by direct sequencing, and length variation between alleles within an individual probably caused the small number of failed sequencing attempts.

3.7. Relative rates of molecular evolution

Likelihood ratio tests were used to investigate rate heterogeneity across the phylogenetic trees. The likelihood was compared between trees with a single rate parameter enforced (strict molecular clock hypothesis) and an alternative hypothesis in which rates were allowed to vary across all branches. In all five tests the strict molecular clock hypothesis was not a significantly worse fit to the data (combined mtDNA data, *Ef1*α, *tektin*, *wg*, and *RpL5*). This implies a relative constancy of evolutionary rates across the tree for all of the gene regions. Nonetheless, as expected evolutionary rates varied considerably between gene regions. Most notably the mitochondrial genes evolve much faster than nuclear genes and rapidly become saturated. At about 1% divergence at *Ef1*α (Fig. 6A) and *tektin* (data not shown) mitochondrial divergence reaches 5–6% and begins to plateau. A similar pattern was seen at *wg*, although there are less variable sites for comparison in this locus (data not shown). In contrast, *RpL5* shows a faster rate of evolution, almost comparable to the mtDNA (Fig. 6B).

3.8. Monophyly of Ithomia

The placement of the root of the mtDNA gene tree was somewhat unstable. When *Napeogenes* was used to root the tree, *Pagyris* fell within *Ithomia*, basal to the *I. agnosia* clade in the phylogenies resulting from either Bayesian or Maximum Parsimony analysis. The placement of *Pagyris* within *Ithomia* did not markedly affect other *Ithomia* relationships. Furthermore, excluding *Napeogenes* and using *Pagyris* to root the mtDNA tree, led to very similar topology and Bayesian posterior probability measures. Trees were also analyzed using a far more distantly related outgroup, *Heliconius* (Beltrán et al., 2002) GenBank Accession No. AF413708 with little change in the composition or relationships of major clades. The instability in the root of the mtDNA tree is most likely due to homoplasy caused by nucleotide saturation. When *Napeogenes* was used to root a maximum parsimony tree inferred using only first codon positions, *Ithomia* was monophyletic with *Pagyris* sister, as expected from morphological characters. Furthermore, *Ithomia* was monophyletic with respect to *Pagyris* when *Napeogenes* was used to root the *Ef1*α, *tektin* and *wg* gene trees. The nuclear gene trees support our assertion that paraphyly of *Ithomia* in the mtDNA phylogenetic analysis was an artifact of nucleotide saturation deep in the tree. The analyses presented below were carried out separately using both *Pagyris* and *Napeogenes* as outgroups with no major differences in the results.

3.9. Comparison of the mtDNA and nuclear DNA trees

Partition homogeneity tests were performed using mtDNA and each of the nuclear genes as data partitions. Homogeneity was rejected ($P = 0.01$) when either *tektin* or *Ef1*α were paired with the mtDNA data partition. The *wg* and mtDNA comparison failed to reject the null hypothesis ($P = 0.84$), although the test was weak owing to the small number of informative sites in the *wg* data. Homogeneity was similarly rejected

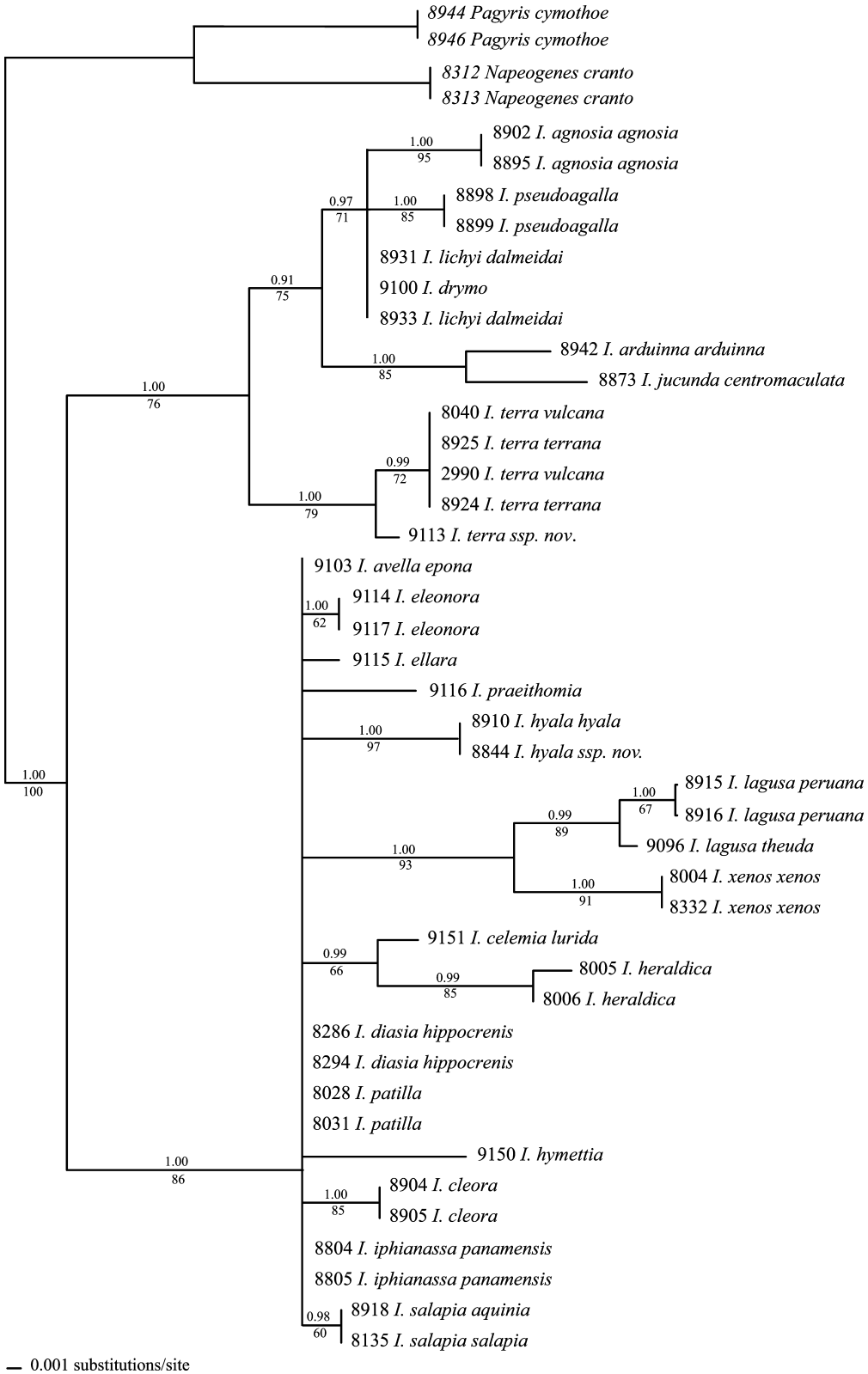

Fig. 4. Phylogenetic hypothesis for *Ithomia* based on 405 nucleotide sites (54 variable) of the *wg* gene. Tree topology was inferred using Bayesian methods. Shown is the consensus phylogeny calculated from trees sampled following the 100,000th generation based on the K80 + G (Kimura, 1980) model. Bayesian probabilities (above) and parsimony bootstrap support (below) is given for each branch.

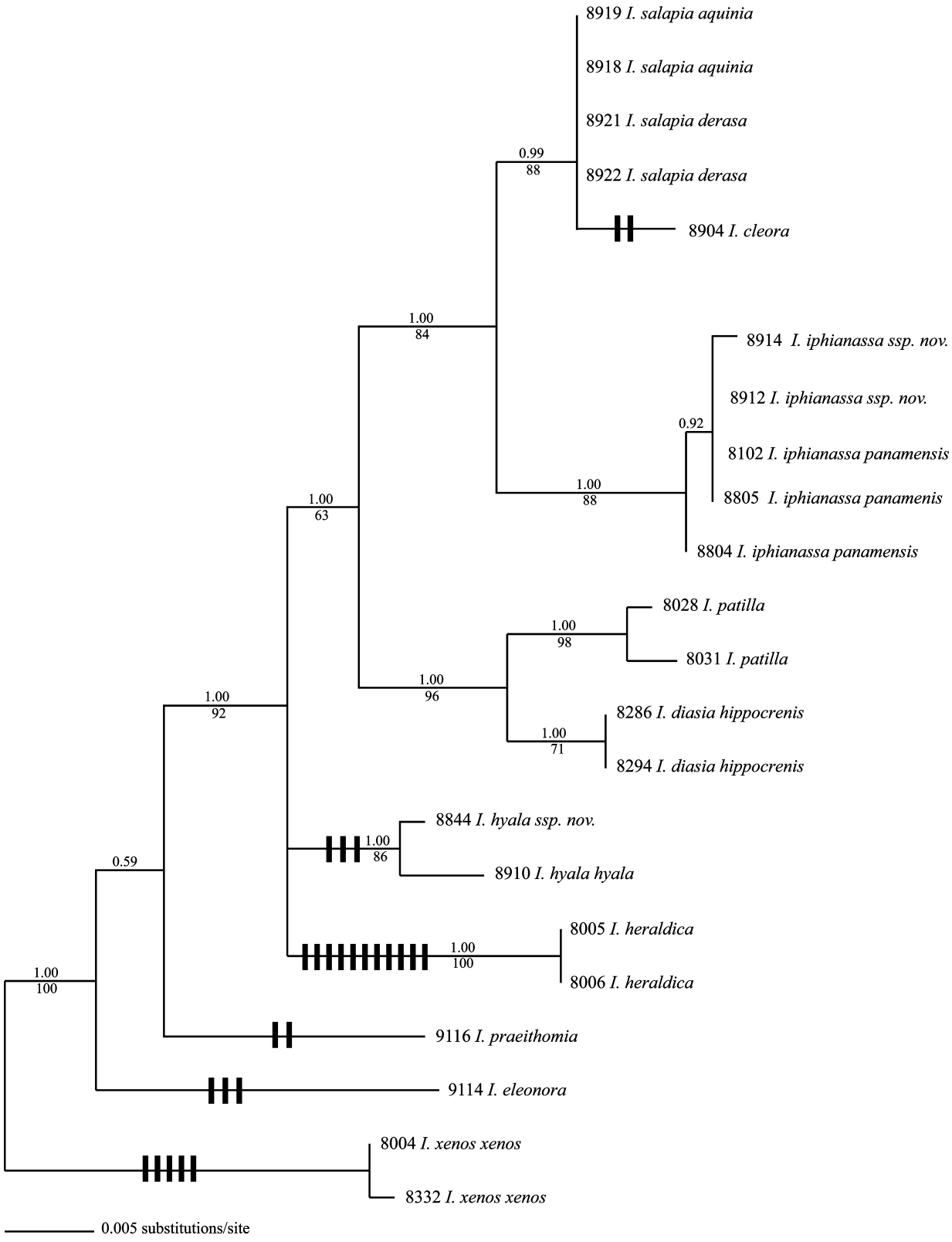

Fig. 5. Phylogenetic hypothesis for *Ithomia* based on 722 nucleotide sites (88 variable) of *RpL5*. Tree topology was inferred using Bayesian methods. Shown is the consensus phylogeny calculated from trees sampled following the 100,000th generation based on the TrN + G model (Tamura and Nei, 1993). Bayesian probabilities (above) and parsimony bootstrap support (below) is given for each branch. Bold vertical lines show synapomorphic indel variation.

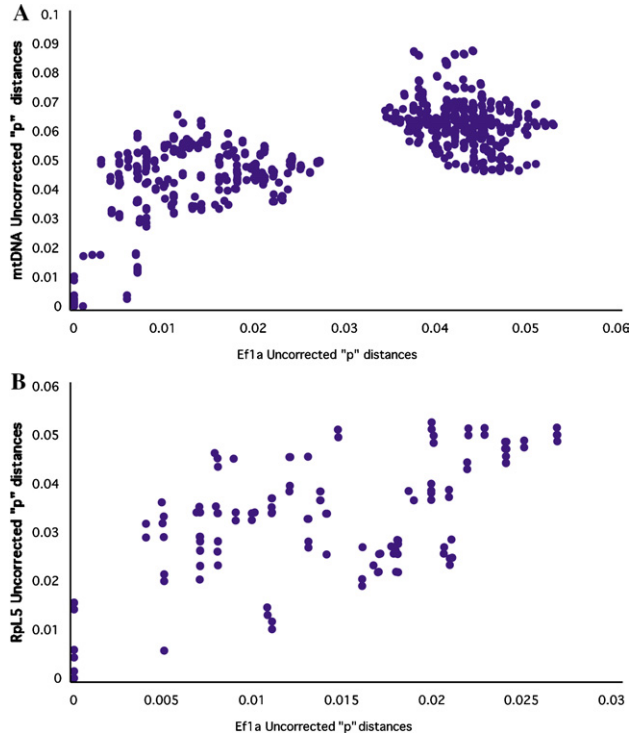

Fig. 6. Plot of relative divergence rates of different *Ithomia* genes. (A) *mtDNA* and (B) *RpL5* uncorrected distances are plotted against *Ef1α* distance for all pairwise species comparisons.

($P = 0.023$) when *Ef1α* was tested with *tektin*. Shimodaira–Hasegawa (SH) tests implemented in PAUP* demonstrated a similar lack of congruence between genes (Table 4).

To investigate the cause of the incongruence among gene trees we used SH tests to compare specific hypotheses regarding the different relationships suggested by the various genes. To do this, topological differences between every possible pair of loci were recorded (i.e., mtDNA vs. *Ef1α*, mtDNA vs. *tektin*, mtDNA vs. *wg*; *Ef1α* vs. mtDNA, *Ef1α* vs. *tektin*, *Ef1α* vs. *wg*; *tektin* vs. mtDNA, *tektin* vs. *Ef1α*, *tektin* vs. *wg*; *wg* vs. mtDNA, *wg* vs. *Ef1α*, and *wg* vs. *tektin*). The topology of the tree generated for each data set was manually altered to mirror specific relationships from another data set, and SH tests were performed to compare the resulting topology against the best tree for a particular gene (Table 5).

In general there was a fairly striking pattern whereby the mtDNA data only weakly rejected the topology suggested by the nuclear data, but the nuclear data sets strongly rejected the mtDNA topology. For example, the nuclear gene trees show *I. terra* as basal to the *I. agnosia* clade, whereas the mtDNA tree grouped *I. terra* as a sister species to the clade formed by *I. agnosia* + *I. lichyi* + *I. pseudoagalla* + *I. drymo* (Fig. 1). SH tests revealed that placing *I. terra* basal to the clade was not a significantly worse fit to the mtDNA data ($P > 0.05$), but *Ef1α* strongly rejected the *I. amarilla* + *I. arduinna* + *I. jucunda* clade as basal to the *I. agnosia* clade ($P < 0.005$). In this case we conclude that *I. terra* basal to the *I. agnosia* clade is the best-supported relationship.

Similarly, *I. xenos* and *I. lagusa* fell within the *I. eleonora* clade in the mtDNA tree while nuclear trees placed them as basal to the *I. eleonora* and *I. iphianassa* clades. As in the previous case, SH tests revealed that the mtDNA data failed to reject the nuclear gene topology ($P = 0.075$) whereas each nuclear gene strongly rejected the mtDNA topology ($P < 0.05$; Table 5). Additional evidence supporting the nuclear gene topology for this particular relationship comes from the high Bayesian posterior probabilities and bootstrap supports of the respective branches in the nuclear trees.

Table 4
Results of SH tests comparing the 'best fit' *Ithomia* tree topologies inferred from different gene regions

Data set				
Topology	MtDNA	Ef1a	Tektin	wg
MtDNA	*	146.21 (0.000*)	88.56 (0.000*)	79.57 (0.000*)
Ef1a	175.29 (0.000*)	*	27.58 (0.023*)	67.34 (0.000*)
Tektin	317.49 (0.000*)	104.11 (0.000*)	*	83.08 (0.000*)
wg	414.74 (0.000*)	229.93 (0.000*)	152.06 (0.000*)	*

Data sets of loci shown vertically were compared against topologies of data sets shown horizontally. The ΔLnl (*P* value) is given for each comparison.

Table 5
SH tests of specific topological hypotheses for *Ithomia* species inferred from different gene regions

Topological differences	MtDNA vs. Ef1a	Ef1a vs. mtDNA	Tektin vs. mtDNA	wg vs. mtDNA
Position of *I. terra*	3.19 (0.209)	21.41 (0.005*)	12.21 (0.054)	1.02 (0.337)
Position of *I. lichyi*	0.10 (0.389)	0.57 (0.383)	4.07 (0.226)	NC
Position of *I. hyala*	5.42 (0.122)	27.53 (0.019*)	27.10 (0.008*)	NC
Position of *I. heraldica* + *I. celemia*	36.34 (0.003*)	4.54 (0.152)	9.94 (0.087)	NC
Position of clade *I. lagusa* + *I. xenos*	12.64 (0.075)	50.62 (0.007*)	37.76 (0.006*)	NC
Sister species of *I. iphianassa* + *I. salapia* clade	23.08 (0.027*)	14.81 (0.039*)	27.78 (0.011*)	NC

In each pair, data sets of the first locus shown were compared against the topology suggested by the second locus. The ΔLnl (*P* value) is given for each comparison. NC, not compared. See text for explanation.

Contrary to the two previous examples, in the case of *I. heraldica* and *I. celemia*, the topology of the mtDNA tree strongly rejected the topology suggested by the nuclear data. In this case *I. heraldica* and *I. celemia* fall within the *I. iphianassa* clade in the nuclear trees but are basal to the *I. iphianassa* clade in the mtDNA tree. SH tests showed that mtDNA strongly reject the nuclear topology and result in a significantly worse fit to the data ($P < 0.005$), but that the nuclear data, both *EF1α* ($P = 0.152$) and *tektin* ($P = 0.087$), failed to reject the mtDNA topology. We therefore conclude that the mtDNA gene topology is the more likely in this instance.

In other cases species relationships were poorly resolved in all data sets. For example none of the possible relationships between *I. agnosia*, *I. pseudoagalla*, *I. drymo*, and *I. lichyi* were a significantly worse fit ($P > 0.05$) to any of the data sets. The lack of resolution is evident from the low bootstrap and Bayesian posterior probability support for these particular branches in the original trees.

One species that does seem to show genuine and strongly supported discordance between data sets is *I. hyala*. While mtDNA, *tektin*, *wg*, and *Rpl5* support the placement of this species with the *I. iphianassa* clade, *Ef1α* shows it basal to the *I. eleonora* and *I. iphianassa* clades. Both of these hypotheses were rejected by at least one data set at $P < 0.05$. It seems likely that this is a genuine case of incongruence between data partitions. However, since four of the five genes are consistent with one another, *Ef1α* is the least likely to reflect the evolutionary history of *I. hyala*.

3.10. The iphianassa/salapia relationship

The tree derived from *RpL5* clearly differentiated the sister species *I. salapia* and *I. iphianassa*, supporting the taxonomic separation of these forms into distinct species (Fig. 5). This result contrasts with the lack of differentiation between these two species in both mtDNA and nuclear exon data. The *RpL5* region also provided additional perspective on the topological disagreements between nuclear and mtDNA data. First, there was clear support for the grouping of *I. cleora* as the closest relative of *I. salapia* + *I. iphianassa*, in agreement with the nuclear exon data. The mtDNA tree shows *I. diasia* as the closest relative of this species pair, whereas all of the nuclear genes show *I. cleora* as the sister species to the pair. After performing the corresponding tests, both mtDNA and nuclear genes rejected the alternate hypotheses ($P < 0.05$). Furthermore, MP bootstrap and Bayesian posterior probabilities are high both in nuclear and mtDNA trees.

3.11. Combined evidence phylogeny

Despite the alternative relationships supported in some parts of the mtDNA and nuclear gene trees, we generated a combined evidence phylogeny derived from concatenated sequences of the six genes, *Co1*, *leucine tRNA*, *Co2*, *Ef1α*, *tektin*, and *wg* (*RpL5* was not included as only a subset of the species were sampled for this locus). A single Bayesian analysis run was carried out using independent models and parameter values for each

Table 6
Partitioned Bremer support analysis for the combined data set

Node number	Total Bremer support	Gene partition					
		Co1	tRNA	Co2	EF1a	Tk	wg
1	16	2.1	0	4.9	3.9	4.2	1
2	7	1.9	0	1.2	0.7	3.2	0
3	28	3.4	0	4.2	14.2	5.2	1
4	7	4	0	4	−0.5	−0.5	0
5	3	−1	0	3	0	−1	2
6	10	3	0	6	1	0	0
7	11	−1	1	−2	2	7	4
8	19	3.4	0	2.2	3.2	8.2	2
9	8	8	0	0	0	0	0
10	0	−2.5	0	−3	2	3	0.5
11	3	−2.35	0	−1.3	2	4.2	0.5
12	12	1.1	0	−0.3	9.2	1	1
13	20	4.2	0	1.2	4.2	8	2.5
14	5	6.5	0	0.3	−1.9	0.1	0
15	5	2.7	0	−4.3	4	3.2	−0.5
16	7	−0.5	0	1	2	3	1.5
17	1	2.2	0	0.7	−2.1	0.2	0
18	7	6.6	0	−0.5	−6.5	7.2	0.2
19	7	0.2	0	0.7	1	4.2	1
20	21	6.4	0	2.9	4.9	4.9	2
Total	197	48	1	21	43	65	19

Node numbers correspond to those shown on Fig. 8.

of the genes. The tree obtained showed very high Bayesian posterior probability and Maximum Parsimony bootstrap support for most relationships (Table 6, Figs. 7 and 8). This combined evidence tree recovered all of the most probable hypotheses determined in the topology comparisons described above. For example, although branch support was weak for the monophyly of the *I. iphianassa* clade relative to *I. hyala*, reflecting the discordance between genes in the position of the latter species, the position of *I. hyala* was concordant with the mtDNA and *tektin* data, rather than the discordant signal from *Ef1α*. In addition, support was weak in parts of the *I. agnosia* clade and at the base of the *I. eleonora* clade, reflecting the lack of resolution across all genes in these two regions of the tree. The strict consensus maximum parsimony tree and the Bayesian tree were completely concordant except in those areas that were poorly supported in both analyses, most notably the position of the species *I. hyala* and *I. hymettia* (compare Figs. 7 and 8).

4. Discussion

4.1. Topological discordance

The final phylogenetic hypothesis for the genus *Ithomia* is based on combined evidence from all gene regions except *RpL5*. This analysis is largely concordant with the relationships deduced from our analysis of individual gene trees. Support values were weak primarily in those regions with poor resolution in all the single gene trees, such as in the *I. agnosia* clade and the base of the *I. eleonora* clade. We also observed weak branch support in regions of the phylogeny where different genes were in disagreement, such as the relationship between *I. hyala* and *I. heraldica*. However, such cases of conflicting phylogenetic relationships between the different gene trees were either not strongly supported in the comparative analyses of alternative topologies, or pitted a single discordant gene or linkage group against all others and were thus easily resolved. We therefore conclude that our combined evidence topology is the best phylogenetic hypothesis consistent with the individual gene trees.

Lack of phylogenetic congruence among different linkage groups representing the same suite of taxa is not surprising (Smith et al., 2004; Sota and Vogler, 2001), and has at least three explanations. First, phylogeny estimation is dependent on a number of assumptions, such as the model of sequence evolution, and is subject to error. Thus, genes might yield different phylogenetic outcomes as a result of erroneous assumptions in the model of nucleotide substitution. Second, coalescence of alleles in isolated populations is a stochastic process and can lead to alleles with different evolutionary histories becoming fixed in different parts of the genome. Third, divergence and speciation is not always the simple bifurcating process typically depicted by phylogenetic trees. Allelic variation can be exchanged between species long after speciation through hybridization and introgression, with the probability of exchange differing markedly between loci, depending on the hitchhiking effect of loci under selection.

The most common pattern of discordance between *Ithomia* gene trees pitted the mtDNA linkage group against the four nuclear gene loci, and may owe to problems modeling the process of mtDNA nucleotide substitution over time. When total uncorrected substitutions in the mtDNA sequences are plotted against the likelihood estimates of divergence derived from the nuclear data (Fig. 6A), saturation in the mtDNA genes is observed around 5–7% divergence, corresponding to 1% divergence in the *Ithomia Ef1α* and *tektin* data. Given that lack of concordance between the mtDNA tree and the nuclear gene trees is most evident at deeper phylogenetic levels, it seems likely that the discordance owes in part to homoplasy in the mtDNA data resulting from multiple substitutions at two and fourfold degenerate codon sites. Nonetheless, a tree derived from mtDNA 1st codon positions had a very similar topology, albeit with weaker branch support, to that from the complete data set (data not shown). This implies that the incongruence with nuclear loci is not entirely due to fast evolving sites, and to some extent probably reflects a true phylogenetic signal.

Gene tree dissonance exhibited by closely related species is less likely an artifact of problems modeling the nucleotide substitution process, and rather suggests that at least one linkage group has an evolutionary history distinct from the species history. *Ithomia diasia* is the closest relative of the *I. salapia* and *I. iphianassa* species pair according to the mtDNA tree, whereas *I. cleora* is the near relative to *I. salapia* and *I. iphianassa* in all nuclear gene trees. The topology difference between the mtDNA and nuclear genes was strongly supported, suggesting that the apparent difference in the evolutionary histories is real. One possible explanation is introgression. The most parsimonious hypothesis would be that the three independent nuclear markers represent the history of relationships among the species, and that the mtDNA history differs owing to introgression. Under this scenario, the pattern of phylogenetic relationships suggests the introgression of *I. salapia* or *I. iphianassa* mtDNA (or that of their ancestor) into *I. diasia* sometime after their split from *I. cleora*. *Ithomia diasia* is sympatric with *I. iphianassa* in Central America and western Ecuador, providing the geographic overlap required for hybridization. Introgression of mtDNA between species is well documented and is often expected to be more likely than at nuclear gene loci when selection is weak or absent owing to the lower effective population size of the mitochondrial genome (Banford et al., 1999; Chow and Kishino, 1995; Rohwer et al., 2001). Introgression of mtDNA may be

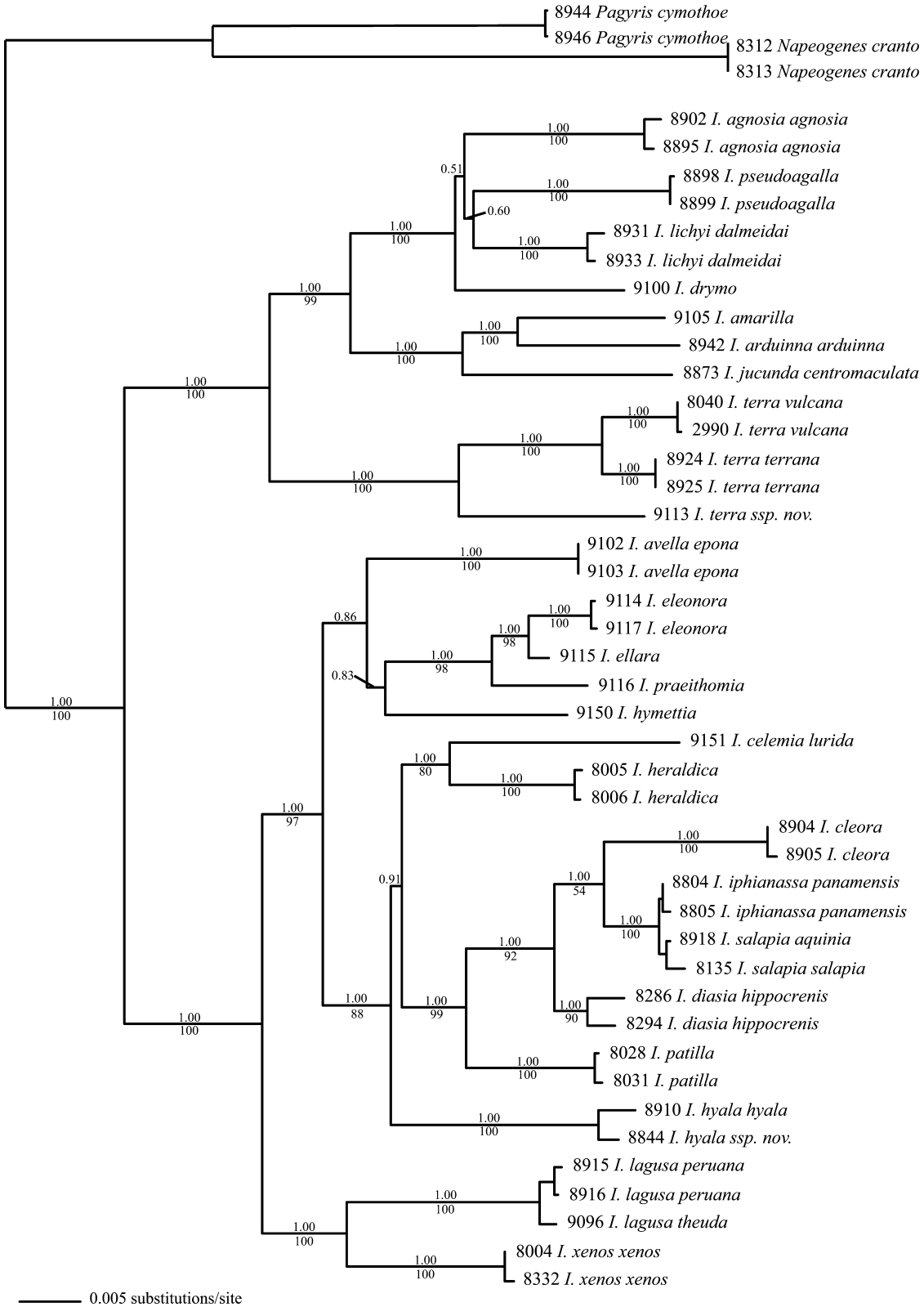

Fig. 7. Phylogenetic hypothesis for *Ithomia* based on sequences of the *Co1*, *leucinetRNA*, *Co2*, it Ef1α, *tektin*, and *wg* genes (3747 nucleotide sites). Genes were combined into a single Bayesian analysis run using independent models and parameter values for each of the genes (see text) Bayesian probabilities (above) and parsimony bootstrap support (below) is given for each branch.

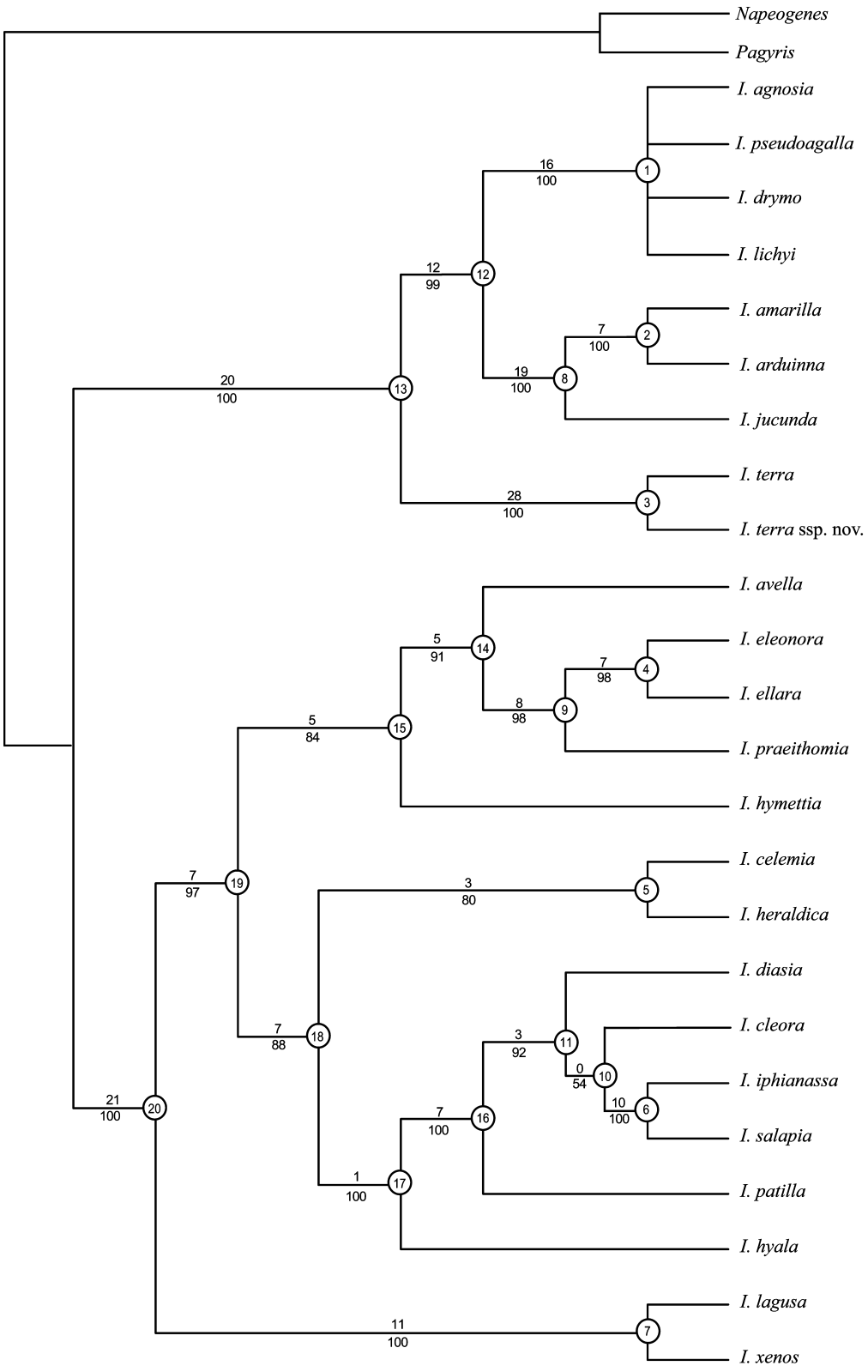

Fig. 8. Maximum parsimony tree for *Ithomia* based on sequences of the *Co1*, *leucinetRNA*, *Co2*, *Ef1α*, *tektin*, and *wg* genes (3747 nucleotide sites). The strict consensus of 10 equally parsimonious trees from the combined data set of 6 gene partitions is shown (length = 1431; CI = 0.633; RI = 0.689). A single individual of each species was selected for this analysis. Bremer (above) and Bootstrap (below) values are shown adjacent to corresponding nodes. Circled node numbers refer to nodes used in calculating total and Partitioned Bremer Support values (Table 6). The MP tree was generated using the branch-swapping algorithm with tree-bisection-reconnection (TBR).

less probable in Lepidoptera, where hybrid breakdown often reduces the viability or fertility of females, but nonetheless cannot be ruled out (Sperling, 1993).

4.2. Taxonomic implications

The *Ithomia* species level phylogeny presented here resolves several uncertainties regarding the evolutionary histories and boundaries that mark closely related lineages in the group. In general genetic divergence between conspecific individuals was low relative to interspecific variation, but there were two notable exceptions. Races of *I. agnosia* and *I. terra* found on the east and west flanks of the Andes exhibit mtDNA divergences of 6.04 and 7.41%, respectively. These genetic distances are roughly an order of magnitude greater than the average genetic distance of 0.54% found between geographic races of other species sampled here, suggesting a moderately long period of separation. In the case of *I. agnosia*, there is no known transition zone between western Ecuadorian and Colombian populations. Indeed, the west Ecuadorian population appears to be geographically isolated, occurring only in moist forests of central western Ecuador, being absent in the wetter forests in the far northwest (Esmeraldas and Carchi provinces—Willmott and Hall, unpublished data). We therefore, consider this west Ecuador population as a distinct species, *Ithomia pseudoagalla* stat. rev. (Rebel). In the case of *I. terra* we retain the present taxonomy, but highlight that the eastern population should probably also be considered as a distinct species if biogeographic evidence corroborates the genetic divergence observed here. The correct nomenclature here is uncertain, as the type specimen of *I. terra* is from Colombia, without more precise data, and may therefore represent either east or west Andean populations. Furthermore, more extensive sampling is needed to determine whether the mtDNA divergence is concordant with morphologically distinct and/or geographically isolated populations.

Also notable from a taxonomic perspective is the large divergence between *I. heraldica* and *I. iphianassa* (7.26% in mtDNA, 1.48% in *tektin*, 0.8% in *Ef1*α, and 1.7% in *wg*), as it has been suggested that these might be geographic races of the same species —this is clearly not the case. Similarly, the form *hymettia* has been classified as a race of *I. lagusa* (Lamas, 2004). However, based on field observations of *I. hymettia* sympatric with *I. lagusa* in Colombia (Sandra Muriel, Pers. Comm.), we considered *I. hymettia* as a distinct species, a conclusion that is clearly confirmed by the molecular analysis (Fig. 7).

The trans-Andean split within *I. terra* and that between *I. agnosia* and *I. pseudoagalla* might represent vicariant divergence associated with the rise of northern Andean cordilleras. However, this is difficult to distinguish from an alternative hypothesis in which occasional dispersal led to the establishment of isolated populations that subsequently diverged. The divergence between these two pairs of taxa is similar, 7.41 and 6.04%, respectively, and suggests divergence times of around 3 million years, assuming 1.1% divergence per lineage per million years (Brower, 1994). However, a recent molecular clock calibration in the genus *Papilio*, based on biogeographic constraints, suggested a much slower rate of molecular evolution (Zakharov et al., 2004) that would give divergence time estimates of around 8 million years for these species. The latter would certainly be consistent with vicariance associated with the rise of the Andes, while the former would suggest trans-Andean dispersal subsequent to the final rise of the northern Andes. Resolution of such issues will require a better calibration of the molecular clock for these butterflies.

In *Ithomia* there is a strong general relationship between phylogeny and geography. The *I. agnosia* clade is restricted almost entirely to southeastern Brazil and the Amazon Basin, the *I. eleonora* clade is restricted to the eastern and southern Andean region, and the *I. iphianassa* clade is found almost exclusively in the western Andes and Mesoamerica. This suggests a pattern of diversification within regions rather than broad-scale vicariance. In addition, many sister species pairs have allopatric or parapatric distributions suggesting a role for geography in speciation. For example, the very closely related species pair *I. iphianassa* and *I salapia* are parapatric and likely come into contact in the eastern slopes of the Colombian Andes. Other species pairs suggest allopatric divergence, such as *I. xenos* and *I. lagusa*, which are distributed in Central America (trans-Andean) and to the east of the Andes (cis-Andean), respectively. Their divergence could have been driven by geological events such as the formation of the Andes mountains and the closure of the Isthmus of Panama, in combination with past fluctuations in vegetation levels due to climate change. The phylogenetic hypothesis presented here will allow future assessment of the roles that geography and wing color have played in the diversification of these butterflies and will contribute to a better understanding of the evolutionary history of *Ithomia*.

Acknowledgments

For the donation of specimens we thank Andrew Brower, Luis Mendoza Cuenca, Gerardo Lamas, Sandra Muriel, and Marie Zimmermann. For discussion and support, we thank Gerardo Lamas, Oris Sanjur, Jesus Mavarez, and James Mallet. For their support both of our collecting trips, and of some fine ithomiine habitat in Panama, we thank Café Duran, EGE Fortuna and the Cana field station. The work was funded by the Smithsonian Tropical Research Institute and the Royal Society. KRW thanks the Leverhulme Trust for support through a postdoctoral fellowship from Standard Research Project Grant F/00696/C.

References

Avise, J.C., 1994. Molecular Markers, Natural History and Evolution. Chapman and Hall, London, pp. xiv + 511.

Banford, H.M., Bermingham, E., Collette, B.B., McCafferty, S.S., 1999. Phylogenetic systematics of the *Scomberomorus regalis* (Teleostei: Scombridae) species group: molecules, morphology and biogeography of Spanish mackerels. Copeia, 596–613.

Barraclough, T.G., Vogler, A.P., 2000. Detecting the geographical pattern of speciation from species-level phylogenies. American Naturalist 155, 419–434.

Barraclough, T.G., Vogler, A.P., Harvey, P.H., 1998. Revealing the factors that promote speciation. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences 353, 241–249.

Bates, H.W., 1862. Contributions to an insect fauna of the Amazon valley. Lepidoptera: Heliconidae. Transactions of the Linnean Society of London 23, 495–566.

Beccaloni, G., 1997. Vertical stratification of ithomiine butterfly (Nymphalidae: Ithomiinae) mimicry complexes: the relationship between adult flight height and larval host-plant height. Biological Journal of the Linnean Society 62, 313–341.

Beltrán, M., Jiggins, C.D., Bull, V., McMillan, W.O., Bermingham, E., Mallet, J., 2002. Phylogenetic discordance at the species boundary: gene genealogies in *Heliconius* butterflies. Molecular Biology and Evolution 19, 2176–2190.

Blum, M.J., Bermingham, E., Dasmahapatra, K., 2003. A molecular phylogeny of the neotropical butterfly genus Anartia (Lepidoptera: Nymphalidae). Molecular Phylogenetics and Evolution 26, 46–55.

Brower, A.V.Z., 1994. Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. Proceedings of the National Academy of Sciences of the United States of America 91, 6491–6495.

Brower, A.V.Z., DeSalle, R., 1998. Patterns of mitochondrial versus nuclear DNA sequence divergence among nymphalid butterflies: the utility of *wingless* as a source of characters for phylogenetic inference. Insect Molecular Biology 7, 73–82.

Brown, K.S., 1984. Adult-obtained pyrrolizidine alkaloids defend ithomiine butterflies against a spider predator. Nature (London) 309, 707–709.

Cho, S.W., Mitchell, A., Regier, J.C., Mitter, C., Poole, R.W., Friedlander, T.P., Zhao, S.W., 1995. A highly conserved nuclear gene for low-level phylogenetics—elongation factor-1-alpha recovers morphology-based tree for Heliothine moths. Molecular Biology and Evolution 12, 650–656.

Chow, S., Kishino, H., 1995. Phylogenetic relationships between tuna species of the genus *Thunnus* (Scombridae: Teleostei): inconsistent implications from morphology, nuclear and mitochondrial genomes. Journal of Molecular Evolution 41, 741–748.

Danforth, B.N., Ji, S.Q., 1998. Elongation factor-1 alpha occurs as two copies in bees: implications for phylogenetic analysis of EF-1 alpha sequences in insects. Molecular Biology and Evolution 15, 225–235.

Davies, N., Bermingham, E., 2002. The historical biogeography of two Caribbean butterflies (Lepidoptera: Heliconiidae) as inferred from genetic variation at multiple loci. Evolution 56, 573–589.

Drummond, B.A., Brown, K.S., 1987. Ithomiinae (Lepidoptera: Nymphalidae): summary of known larval food plants. Annals of the Missouri Botanical Garden 74, 341–358.

Farris, J.S., Källersjö, M., Kluge, A.C., Bult, C., 1995. Constructing a significance test for incongruence. Systematic Biology 44, 570–572.

Fox, R.M., 1940. A generic review of the Ithomiinae (Lepidoptera: Nymphalidae). Transactions of the American Entomological Society 66, 161–207.

Fox, R.M., 1968. Ithomiidae (Lepidoptera: Nymphaloidea) of Central America. Transactions of the American Entomological Society 94, 155–208.

Huelsenbeck, J.P., Ronquist, F., Nielsen, R., Bollback, J.P., 2001. Evolution—Bayesian inference of phylogeny and its impact on evolutionary biology. Science 294, 2310–2314.

Jiggins, C.D., Estrada, C., Rodrigues, A., 2004a. Mimicry and the evolution of pre-mating isolation in *Heliconius melpomene*. Journal of Evolutionary Biology 17, 680–691.

Jiggins, C.D., Mavarez, J., Beltrán, M., Johnston, J.S., Bermingham, E., 2004. A genetic map of the mimetic butterfly, *Heliconius melpomene*. Genetics, in press.

Jiggins, C.D., Naisbit, R.E., Coe, R.L., Mallet, J., 2001. Reproductive isolation caused by colour pattern mimicry. Nature (London) 411, 302–305.

Kimura, M., 1980. A simple method for estimating evolutionary of base substitution through comparative studies of nucleotide sequences. Journal of Molecular Evolution 16, 111–120.

Lamas, G., 2004. Ithomiinae. In: Heppner, J.B. (Ed.), Atlas of Neotropical Lepidoptera. Checklist: Part 4A. Hesperioidea—Papilionoidea. Association for Tropical Lepidoptera/Scientific Publishers, Gainsville.

Lerat, E., Daubin, V., Moran, N.A., 2003. From gene trees to organismal phylogeny in prokaryotes: the case of the gamma-proteobacteria. PLoS Biology 1, 101–109.

Maddison, W.P., Maddison, D.R., 1997. MacClade: Analysis of Phylogeny and Character Evolution. Sinauer Associates, Sunderland, MA.

Mita, K., Morimyo, M., Okano, K., Koike, Y., Nohata, J., Kawasaki, H., Kadono-Okuda, K., Yamamoto, K., Suzuki, M.G., Shimada, T., Goldsmith, M.R., Maeda, S., 2003. The construction of an EST database for *Bombyx mori* and its application. Proceedings of the National Academy of Sciences of the United States of America 100, 14121–14126.

Pääbo, S., 2003. The mosaic that is our genome. Nature (London) 421, 409–412.

Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Rohwer, S., Bermingham, E., Wood, C., 2001. Plumage and mitochondrial DNA haplotype variation across a moving hybrid zone. Evolution 55, 405–422.

Rokas, A., Williams, B.L., King, N., Carroll, S.B., 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. Nature (London) 425, 798–804.

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., Flook, P., 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. Annals of the Entomological Society of America 87, 651–702.

Smith, V.S., Page, R.D.M., Johnson, K.P., 2004. Data incongruence and the problem of avian louse phylogeny. Zoologica Scripta 33, 239–259.

Sorenson, M.D., 1999. TreeRot. Boston University, Boston, MA.

Sota, T., Vogler, A.P., 2001. Incongruence of mitochondrial and nuclear gene trees in the carabid beetles *Ohomopterus*. Systematic Biology 50, 39–59.

Sperling, F.A.H., 1993. Mitochondrial-Dna variation and Haldane rule in the Papilio-Glaucus and P-Troilus species groups. Heredity 71, 227–233.

Sperling, F.A.H., Harrison, R.G., 1994. Mitochondrial DNA variation within and between species of the *Papilio machaon* group of swallowtails. Evolution 48, 408–422.

Swofford, D.L., 2000. PAUP* Phylogenetic Analysis Using Parsimony (* and Other Methods). Sinauer Associates, Sunderland, MA.

Tamura, K., Nei, M., 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Molecular Biology and Evolution 10, 512–526.

Tobler, A., Kapan, D.D., Flanagan, N.S., Gonzalez, C., Peterson, E., Jiggins, C.D., Johnston, J.S., Heckel, D.G., McMillan, W.O., 2004. First generation linkage map of the warningly colored butterfly *Heliconius erato*. Heredity, in press.

Vane-Wright, R.I., 1979. Towards a theory of the evolution of butterfly colour patterns under directional and disruptive selection. Biological Journal of the Linnean Society 11, 141–152.

Wahlberg, N., Weingartner, E., Nylin, S., 2003. Towards a better understanding of the higher systematics of Nymphalidae (Lepidoptera: Papilionoidea). Molecular Phylogenetics and Evolution 28, 473–484.

Yang, Z., 1994. Estimating the pattern of nucleotide substitution. Journal of Molecular Evolution 39, 105–111.

Zakharov, E.V., Caterino, M.S., Sperling, F.A.H., 2004. Molecular phylogeny, historical biogeography, and divergence time estimates for swallowtail butterflies of the genus *Papilio* (Lepidoptera: Papilionidae). Systematic Biology 53, 193–215.